PREFACE

In this thesis, we consider the problem of regularized loss minimization problem of machine learning that can be framed as a non-smooth composite convex loss minimization problem. Applications that adopt this framework include various recognition problems, recommender systems, machine vision, information retrieval, bio-informatics, neuroscience, finance, web mining, medicine and many more. A popular class of techniques that are applied to solve such frameworks is the first-order proximal methods, and a recent trend is to design new accelerated proximal algorithms to solve such frameworks efficiently. Proximal methods can be generalized as the operator splitting techniques, which can be interpreted as the fixed-point iterative schemes. In order to propose and analyze new accelerated gradient algorithms, one approach is to use novel concepts of fixed-point theory. In this work, we have used few concepts of fixed-point theory and designed new accelerated gradient algorithms. These proposed algorithms are then applied to solve various real-world machine-learning problems.

We begin with the concept of extragradients of fixed-point theory. New proximal gradient algorithms (or forward-backward splitting algorithms) based on this concept are designed and a novel extragradient-based accelerated proximal algorithm is proposed. The convergence guarantee and the stability analysis is also given. The practical performances of all the algorithms are tested on three application problems, which are a) the high-dimensional regression task with traditional lasso framework, b) unified sparse representation learning, and c) solving the lasso framework with overlapping group and fused penalties. Several publicly available benchmark real datasets are used in the experimentation for each application. Experimental results indicate that not only the number of iterations to reach the convergence is significantly lesser than the number of iterations consumed by previous algorithms, the error obtained with our algorithm is also significantly lesser than the error we obtained with previous algorithms. These results are consistently better in all the considered applications, which directly indicate the suitability of the proposed algorithm to solve such problems.

The second contribution of this thesis is to introduce the novel concept of viscosity-approximation of fixed-point theory to the field of machine learning. A recently proposed viscosity-approximation based forward-backward splitting technique is designed as a proximal gradient algorithm. In addition, we proposed a novel viscosity-approximation based accelerated gradient algorithm and proved the convergence of the algorithm. The major part of this contribution is that we proved the strong convergence guarantee of the proposed algorithm in the general infinite-dimensional Hilbert spaces. Both the algorithms are applied to solve the regularized multitask learning problems, where the sparsity-induced regularizers are used to frame the relatedness among tasks. The first application is to solve the problem of the multitask regression problem and the second application is the joint splice-site recognition task, which is a popular problem in the field of bioinformatics. For both the applications, we performed experiments on several publicly available real benchmark datasets. We found that solving the multitask framework with the help of viscosity-approximation based proximal algorithms is not only faster, but also these algorithms give a stable solution in comparison to the traditional proximal algorithms.

Our third contribution belongs to the general class of the operator splitting techniques, which relaxes the differentiability assumption we considered in our last two contributions. Here, an extragradient-based operator splitting technique is proposed, which is based on the extragradient fixed-point schemes. An accelerated variant of this algorithm is also proposed, which combines the concept of extragradient fixed-point schemes with the inertial step for the general operator splitting techniques. The convergences of both the algorithms are proved. Both the algorithms are applied to analyze the microarray gene expression datasets for the task of cancer prediction. Four microarray gene expression datasets were used in our experiments. Experimental results indicate that the number of iterations our algorithm takes to reach the convergence is significantly lesser in comparison to the number of iterations consumed by previous algorithms.